Originally Published in: ReNews, March, 2004 PUB-RR-716

MANN-KENDALL TEST - HOW TO APPLY RESULTS

The Mann-Kendall, non-parametric statistical test is rou-tinely used to assess trends in groundwater concentration data. Consultants and other parties working on contami-nated properties often use the Mann-Kendall test and sub-mit their results to the RR Program staff for review. However, program staff have noticed that, in some cases, individuals are applying the Mann-Kendall test results incorrectly. The Mann-Kendall test is required for sites with petroleum contamination above groundwater standards, when par-ties request closure under Comm 46 and Chapter NR 746, Wis. Adm. Code (and under the applicable sections in NR 726, Wis. Adm. Code, that are cross-referenced in Comm46/NR 746). An alternative non-parametric test, Mann-Whitney, is also acceptable.

Please note that closure under Chapter NR 726 does not require any specific statistical test for data assessment. Nevertheless, Mann-Kendall is often used to support closure requests under NR 726 because it is readily available, easy to use, and the DNR has developed tools for its application.

Selecting Appropriate Tests Selection of an appropriate statistical test requires care-ful consideration of the data available and the limita-tions of the test. The Mann-Kendall test procedure starts by simply comparing the most recent round of water-quality datum with the results of all earlier rounds. A score of +1 is awarded if the most recent concentration is larger, or a score of -1 is awarded if it is smaller. The total score for the time-series data is the Mann-Kendall statistic, which is then compared to a critical value, to test whether the trend in concentration is increasing, decreasing or if no trend in concentration can be determined. The simple procedure is illustrated on page 63 of the RR program's Natural Attenuation guidance (publication #RR-614), which is available at the following link: http://dnr.wi.gov/org/aw/rr/archives/pubs/RR614.pdf.

Mann-Kendall Limitations The limitations of the Mann-Kendall (M-K) test include:

Originally Published in: ReNews, June, 2003 PUB-RR-713

1. The test does not account for the magnitude of the data. The test can be applied to a data series in the 10's of ug/l as easily as a data series in the 1000's of ug/l. However, the larger the actual numbers, the less sensitive the test is to changes in the magnitude of numbers. For example, at an 80 percent confidence level, given benzene concentrations of 6, 10, 8, 13, and 10 ug/l, the M-K test would conclude that there was an increasing trend; while a series like 10; 9,000; 8,500; 9,500; and 8,900, the M-K test will conclude that this is a stable trend. In fact, the first series likely represents a stable trend while the second series - with a jump of more than two orders of magnitude - represents an increasing trend.

2. The test does not account for the temporal variation in the data such that we cannot obtain a degradation rate, nor account for variable "releases." Estimating degradation rates is important to determining the expected life time of the contaminant source and plume. To determine degredation rates, other data analysis tools, such as regression analysis, is needed. Actual data values may vary greatly. For instance, when "slugs" of contaminant are released due to seasonal water table fluctuations. Another example - M-K will determine that the following series has a decreasing trend: 0.23; 5; 43; 921; 1,340; 103; 1.62; 0.23; 0.23; and 0.23. In fact, the set of data is more indicative of a contaminant slug moving past the monitoring well (i.e., an advancing plume margin), especially if a more downgradient well shows an increasing trend. When there is not a downgradient well, the

confidence in the M-K result decreases.

3. The data must be free of "seasonality." The underlying assumption for the M-K test is that declining or stable trends are due to the effects of natural attenuation processes. If factors other than NA processes - like differing analytical meth-ods, over- or under-purging wells, using bailers or low-flow pumps - are observed to affect the sample results, use of the M-K test or use of the M-K test results would be suspect, and the test should not even be performed. With the M-K test, the goal is to assess changes in water quality due to **time**. Other fac-tors affect water quality over time, however - seasonality being one of them. Seasonal effects on water quality can be brought about by changes in the water-table elevations or in the groundwater flow gradient and direction. Appendix A of Comm 46 and NR 746 provides methods to determine whether data are affected by seasonality. The RR Program also provides a tool to quickly assess the relationship between water elevation and contaminant concentration. The tool is a spreadsheet available on the program's web site as a zipped file, located at the following link: http://www.dnr.state.wi.us/org/aw/rr/ archives/pub_index.html#10 (RR-614.zip). When the data appear to be seasonally affected, one way to reduce the seasonal bias in the M-K test is to include only data collected from the particular season with the highest contaminant concentrations. This would require a longer period of monitoring to collect sufficient data.

4. A "no trend" result does not equate to a stable plume. The no-trend result simply means that the M-K test could not discern either an up or a down trend for the given set of data, and in fact a "decreasing" or "increasing" result from the M-K test is a more robust conclusion than the "no trend" result.

A complication is brought about because, when the M-K test result shows "no trend," the DNR spreadsheet proceeds to estimate the coefficient of variation (CV) of the concentrations. The spreadsheet accepts a CV of less than or equal to 1 as equivalent to a stable plume. The CV estimation would tend to yield different conclusions for the high concentrations than for low concentrations, and would favor tagging higher concentrations as stable.

This can be easily shown by comparing two "no trend" results: a set of high numbers (e.g. 4,300; 5,500; 8,000; 5,500) and a set of low numbers (85, 393, 22, 45). The absolute difference among the larger-number set is in the 1000's, while among the smaller numbers, at least an order of magnitude less. However, the CV for the larger-number set (CV = 0.3) is less than for the smaller number sets (1.3). Therefore, "stability," when defined as being CV≤1, would favor the larger-number set.

The CV estimation can provide the wrong impression when the M-K test has already failed to discern a trend. While the DNR spreadsheet accepts a no-trend result at an 80 percent confidence level, and a CV≤1 as equivalent to a "stable" plume, the fact is that careful thought must be given to the data before this conclusion can be accepted. This is especially true when a small data set - the minimum of four rounds - is used in the spreadsheet.

It is strongly recommended that more than the minimum four rounds of data be assessed in the M-K test. The department's guidance recommends that at least six rounds of data be used in the spreadsheet. It is only common sense that, the less data available, the less reliable the conclusions will be based on that data.

In a related concern, sometimes a large data set exists but it is truncated in the M-K analysis. That is, the user selects only the last four rounds of data for analysis, even though more data exists. This is inappropriate. The DNR spreadsheet has a limit of 10 data/well. This is because a different M-K algorithm is used for more than 10 rounds of data. If more than 10 rounds of data exist, we recommend the user choose the most appropriate of the following responses:

• assessing the seasonality of the data and test data from the season of the year that exhibits the highest contaminant concentrations;
• use the latest 10 rounds of data; or
• use a different statistical test that will accommodate more data rounds.

The Mann-Kendall test is a tool that is appropriate to use in some situations with some data sets. It should be used only after careful assessment of the data and a determination that the test is appropriate for the data set. Consultants and other parties utilizing the test are encouraged to use the whole range of data assessment tools available. Several alternatives to the Mann-Kendall test for data trends are discussed in the DNR guidance about the test, located at the following link: http://dnr.wi.gov/org/aw/rr/archives/pub_index.html#15. For more information, please contact Terry Evanson at 608-266-0941, or theresa.evanson@dnr.state.wi.us.

AVOID PROBLEMS WITH LABORATORY DATA

In submitting data to the RR Program, individuals must use a certified or registered laboratory per requirements in Chapter NR 149, Wis. Adm. Code (Laboratory Certification and Registration). Some consulting firms have been overlooking this requirement - often as early as during a bid process, but also when submitting work plans, site investigations, remedial activities, during long-term monitoring and sometimes even at requests for case closure.

It is possible that familiarity with a lab can lead to overlooking its current certification status. Not all labs are certified in all test categories. Some categories include multiple parameters, and labs are certified separately for each parameter within the category. In addition, a large analytical service may have multiple locations, and each location needs to be certified for the test categories being analyzed.

Finally, don't forget that certifications need to be renewed. A lab that was certified a year ago for all parameters needed for a site investigation may, for a number of reasons, not be currently certified. It is not unusual for a certified lab to lose or withdraw certification for one or more parameters staff formerly maintained.

Saving Time And Money
The remedy is simple. It's a lot easier, time and money-wise, to avoid lab certification problems by careful planning early in the process, rather than be faced later on with unusable data. After the analytical parameters included in field work have been established, use the following "Commercial Laboratories" link to determine a lab's certification status for those parameters: http://www.dnr.state.wi.us/org/es/science/lc/search/.

Note that other lists on the same web page show the status of labs performing specific analyses (e.g. pesticides, petroleum, etc). Be sure you check the lab location that you will

actually be using. Also, please remember when a certified lab subcontracts with another lab, the subcontracted lab must have the same certification.

For some multi-analyte methods, typically organics, certification is offered under more than one test category, and it's usually based on technology. Be certain that the laboratory holds certification for the technology being used. For example, if the parameter is a polycyclic aromatic hydrocarbon (PAH), HPLC method 8310 is often used due to its low detection limits (test category 13).

However, labs are developing methods for PAHs that operate the mass spectrometer in selected ion monitoring mode, or SIM. This allows detection limits that equal or surpass HPLC. Because SIM is an option under method 8270C, it meets the methods requirements of Chapter NR 716, Wis. Adm. Code. In this case, however, the lab would need to be certified either for semi-volatiles by gas chromatography/mass spectrometry (GC/MS) under test category 12 (organic; semi-volatiles by GC/MS) or for PAHs by GC/MS SIM under test category 19 (any single analyte or analyte group). If a lab can achieve low enough detection limits for PAHs, either method may be used for groundwater analysis.

For More Information If you have questions about laboratory certification, please contact Rick Mealy, DNR Science Services, at 608-264-6006, or richard.mealy@dnr.state.wi.us. For questions about RR Program chemistry issues, please contact Charlene Khazae at 608-267-0543, or charlene.khazae@dnr.state.wi.us.